

Título: El Impacto de FRBR en Argentina: Implementación de un modelo de objetos basado en FRBR, CRM y FRBRoo en CAICYT-CONICET.

Autor: Norberto Manzanos

Correo electrónico: nmanzanos@caicyt.gov.ar; nmanzanos@gmail.com

Institución: CAICYT-CONICET

Dirección: Saavedra 15

Ciudad: Ciudad Autónoma de Buenos Aires

País: Argentina

Resumen: Este artículo presenta un diseño informático para el registro de material documental basado en los modelos FRBR_{ER}, CIDOC CRM y FRBR_{OO} que está siendo desarrollado en CAICYT-CONICET. La propuesta parte del primer grupo de entidades definido por el modelo FRBR_{ER} (*Obra, Expresión, Manifestación e Item*) al que incorpora alguna de las propuestas de los otros dos modelos. En su planteo general, rescata de éstos la utilización del paradigma de objetos, lo que deriva en una definición de conceptos más rigurosa. Toma de ellos, entre otras cosas, la modelización de eventos, lo que permite representar a los documentos en su devenir temporal. Trata de profundizar en otros aspectos que aún no han sido tratados por FRBR_{OO}: la falta de clasificación exhaustiva, el estatus ontológico del *Item* y su relación con el soporte físico, el problema de la responsabilidad y de la función del responsable documental y el problema de los nombres. Dado que el producto presentado está actualmente en uso en CAICYT-CONICET en la base de datos de ISSN argentino, se detalla cual es la aproximación del modelo al problema de las series. Finalmente, se mencionan algunas particularidades de la implementación: uso del lenguaje Smalltalk, su dialecto de código abierto, Squeak, y los frameworks independientes que se han desarrollado: Atón y Smallfaces.

Presentación

El programa Opus tiene como antecedente un trabajo de investigación llevado a cabo durante 2003 junto con los compañeros del CAICYT Tatiana Carsen, Hugo García y Cecilia Mabragna. A causa de los clásicos vaivenes políticos a que está sujeta la administración pública en los países periféricos, el proyecto estuvo detenido poco más de un año. Finalmente, a partir de la gestión de Mario Albornoz, el proyecto ha podido salir de su etapa de investigación y hoy ya es un producto informático en uso. Integran el equipo de desarrollo, además del autor del presente artículo, Juan Matías Burella y Hernán Morales. A todos los nombrados, a Elsa Barber y a las demás autoridades y colegas de la Biblioteca Nacional queremos agradecer el hecho de estar aquí exponiendo nuestro trabajo.

Antes de comenzar es necesario hacer algunas aclaraciones formales. Dado que este modelo comenzó a ser pensado en el 2003 siguiendo la metodología de diseño con objetos -un año antes del modelo CRM y cuatro años antes de la primera versión de FRBR orientada a objetos-, algunos de los aspectos detallados en este artículo fueron desarrollados en forma paralela y se arribó a conclusiones equivalentes. En ocasiones el modelo sufrió influencias que obligaron a tomar otros rumbos, mientras que en otros casos hemos optado por soluciones diferentes y es en este último aspecto en donde más vamos a detenernos. También es necesario aclarar que la mayoría de los ejemplos que damos están simplificados con fines expositivos, pues no es posible mostrar en todo

detalle el modelo dentro de los límites de esta ponencia.

Debemos hacer también una pequeña aclaración con respecto a la terminología. Cuando mencionamos *entidades*, *relaciones*, *clases* o *atributos* definidos en alguno de los tres modelos analizados, conservamos los nombres en su forma original. No obstante, si se encuentran dentro del flujo del texto, utilizamos las traducciones usuales -si las hay- o nuestras, en caso contrario. En cuanto a los nombres técnicos de nuestro modelo, éstos siguen la norma de denominación de clases y variables propias del lenguaje Smalltalk, el usado en el desarrollo: los nombres de clase, en mayúsculas sin espacio; los nombres de mensajes o variables, en minúsculas. El material gráfico consiste en diagramas de clases UML simplificados: los rectángulos identifican clases, que pueden contener nombres de atributos en rectángulos internos; las flechas vacías representan relaciones de herencia, mientras que las flechas llenas representan relaciones de colaboración.

Premisas

Antes de presentar los aspectos teóricos y prácticos de nuestra propuesta nos parece ineludible establecer cuáles han sido las premisas de las cuales hemos partido. Cada una amerita por sí sola una profundización que por cuestiones de espacio nos es imposible abordar en este momento. No obstante, daremos al menos unas justificaciones mínimas.

La ciencia de la documentación ha dado un salto cuantitativo a partir del modelo FRBR, de la mera continuación de técnicas con un marco teórico débil al primer indicio del establecimiento de un modelo conceptual a partir del cual construir una teoría y una práctica consistentes. Iniciativas como CRM y su consecuencia en el mundo de la bibliotecología, el FRBR-_{OO}, demuestran la tendencia a encontrar conceptos comunes en áreas del conocimiento que comparten no tanto técnicas como objetos de estudio: la bibliotecología y la museología. (Le Boeuf 2003a). Por esta razón nuestro punto de partida han sido las definiciones del grupo de entidades 1 de la propuesta FRBR-_{ER}. Mucha literatura ha demostrado que estos conceptos ya existían en los trabajos de Luvetzky (Yee 2000) y otros, siempre dentro del mundo de la ciencia de la documentación o de la bibliotecología. También se ha insistido en que estos conceptos están presentes en forma implícita no sólo entre los especialistas sino también entre los usuarios, como lo evidencia Le Boeuf al analizar los múltiples significados que atribuimos a la palabra "libro" (Le Boeuf 2003b).

Pero también pueden ser rastreados en otros ámbitos. Desde el estructuralismo y la cibernética, Abraham Moles esquematizaba el canal de difusión del pensamiento que constituye la edición de libros de la siguiente forma: "La creación de ideas por el autor (...) [*Obra*], la puesta en forma, que es la tarea de redacción propiamente dicha [*Expresión*], la normalización [por medio de dactilógrafos] que destruye la unicidad del manuscrito, la aceptación de un sistema de difusión [como] el libro, la fabricación de copias, por medio de la impresión [*Manifestación*], la asimilación por parte de los consumidores individuales [*Ítem*]" (Moles 1971).

Teniendo en cuenta que nuestra tarea es una confluencia de la ciencia de la documentación y la informática, y que dentro de ésta última optamos por el paradigma de objetos, debemos justificar mínimamente esta elección, a sabiendas de que es también un tema polémico y que amerita una discusión aparte. Más allá de las consideraciones técnico-económicas (si su uso deriva en

productos más estables, diseños más legibles, sustentables, escalables, etc) y de las de índole práctica (si produce diseños más comunicables y expresivos) consideramos que el problema a tratar es del orden de la representación y que el paradigma de objetos permite una mejor representación de la realidad, tal como la entendemos los humanos. "Un sistema de computación debe proveer modelos que sean compatibles con los de la mente" (Ingalls 1981). "El software tradicional (...) se concentra primera en la función -la función de un programa (...) Pero el conocimiento humano a menudo trabaja de otra forma, reconociendo primero las cosas" (Liu 1996). "En un nivel, la programación con objetos es más natural porque nos permite organizar la información en modos que nos son familiares, como lo ilustra la jerarquía de clases. En un nivel más profundo, es más natural pues refleja las técnicas propias de la naturaleza para abordar la complejidad" (Taylor 1997)

La programación tradicional, por provenir de la matemática, tiende a ser más reduccionista cuanto más complejos son los dominios que trata de representar. El paradigma de objetos, en cambio, parte de metáforas biológicas y lingüísticas, lo que disminuye el reduccionismo en la medida en que nos acercamos más a la forma que tenemos de concebir la realidad.

El aspecto lingüístico-semiótico está presente en esta primera definición del creador de Smalltalk, Alan Kay, cuando definió el lenguaje de programación Smalltalk como un "medio de comunicación" (XEROX 1976). "Programar implica un proceso orientado a las computadoras que acarrea un control rígido, estricto, exacto y lineal, mientras que 'comunicación' es un proceso orientado a los humanos que implica comprensión y consenso, y a menudo es inexacto". Los dos elementos fundamentales de la orientación a objetos, objeto-mensaje, evidencian en forma más clara lo dicho: "Todas las acciones en Smalltalk son realizadas enviando mensajes a objetos, y todo mensaje devuelve un objeto" (Mortensen 2001).

Con respecto a la influencia de la biología, Kay nos dice: "Desde el punto de vista filosófico, los objetos de Smalltalk tienen mucho en común con las mónadas de Leibnitz y las nociones de la física y la biología del siglo XX" (Kay 1993). La metáfora biológica, además, está todo el tiempo presente en términos técnicos claves como *herencia*, *clasificación* y *comportamiento*.

Debemos, por tanto, dar algunas definiciones de la terminología básica de objetos. El paradigma de objetos tiene tres características clave: objetos, mensajes y clases. Un objeto informático está formado por algunas operaciones y alguna información, lo que modifica la concepción estructural, según la cual los programas se construyen con estructuras de datos y algoritmos. Los objetos se comunican entre sí enviándose mensajes, y la respuesta a un mensaje es otro objeto. La herencia es el mecanismo que permite a una clase compartir la estructura y los comportamientos definidos de otra u otras clases. Las clases son "fábricas" de objetos (Liu 1996).

Dado que los modelos CRM y FRBR-_{oo} la utilizan, no podemos dejar de mencionar una cuestión polémica dentro del paradigma de objetos: la herencia múltiple. Los casos de herencia múltiple en dichos modelos son bastante complejos como para utilizar a modo de ejemplo y la conveniencia de su aplicabilidad demandaría al menos un artículo específico, por lo que ilustraremos el problema con ejemplos sencillos. Si queremos representar la idea "los murciélagos, las palomas y los aviones vuelan, mientras que las gallinas y los camiones no" dentro de una jerarquía ya

establecida, tal que "las palomas y las gallinas son aves, los murciélagos son mamíferos y los camiones y los aviones son vehículos", utilizando herencia múltiple podríamos establecer una entidad abstracta "Objeto Volador", de la que algunos de los objetos en cuestión heredarían sus propiedades y otros no. En cambio, utilizando herencia simple, los que tienen la propiedad "volar" simplemente la implementan, sin modificar la jerarquía original. Unos saben responder al mensaje "volar" y otros no. Es lo que se conoce en terminología de objetos como *polimorfismo*: la capacidad de los objetos de distintas clases de tener un mismo comportamiento manifestado de diferentes maneras.

Nuestra posición no es rígida al respecto, pero tendemos a pensar que en el momento en que es necesario que una clase herede características propias de más de una clase, la herencia en tanto herramienta conceptual deja de ser la apropiada y se introduce una complejidad innecesaria: estamos ante los mentados límites de la clasificación y cabe aplicar aquí la vieja enseñanza de Guillermo de Ockham: "Los entes no deben ser multiplicados sin necesidad".

Antes de abordar el tratamiento del modelo propuesto, es necesario hacer una breve referencia a los modos de representación de la información bibliográfica mediante computadoras.

Analizaremos los dos modelos tradicionales y la diferencia que introduce la orientación a objetos. Los llamados SIR (IRS) (Sistemas de Recuperación de información), como observa Moya (Moya Anegón 1995), adolecen de una falta de base teórica en comparación con los otros modelos de información y su justificación ha sido siempre de orden práctico. No obstante, para los fines de este trabajo podemos destacar que la característica principal de estos sistemas desde el punto de vista de la representación es que la entidad o concepto que se quiere representar y su correlato informático son equivalentes. El modelo define una entidad de nivel más elevado (el documento) al cual otras entidades están subordinadas (el autor, el título). Estas entidades secundarias no poseen un correlato informático propio debido a las limitaciones de los sistemas de archivos utilizados, sino que consisten en cadenas de caracteres repetidas. Esta concepción se podría resumir con la ecuación **un documento = un registro**. Si queremos representar la relación "un documento es producido por varias personas" debemos repetir la cadena de caracteres que identifica al autor en cada documento. El problema que trae aparejado esta repetición es, por un lado, la falta de normalización, pues habrá tantas cadenas de caracteres representando a un autor como documentos haya producido éste, pero, por otro, la entidad "autor" no tiene una equivalente físico unívoco. Los registros de autoridades vienen a suplir este problema: se define una tabla de autores y luego se establece el vínculo con los documentos. Pero si queremos representar la relación inversa "una persona puede producir varios documentos", debemos repetir el nombre del documento en cada uno de los autores. En lugar de nombres, podemos usar identificadores más precisos, pero la asignación de éstos, dado el sistema de archivos usualmente utilizado, deben ser asignados manualmente y por ende están sujetos al error humano.

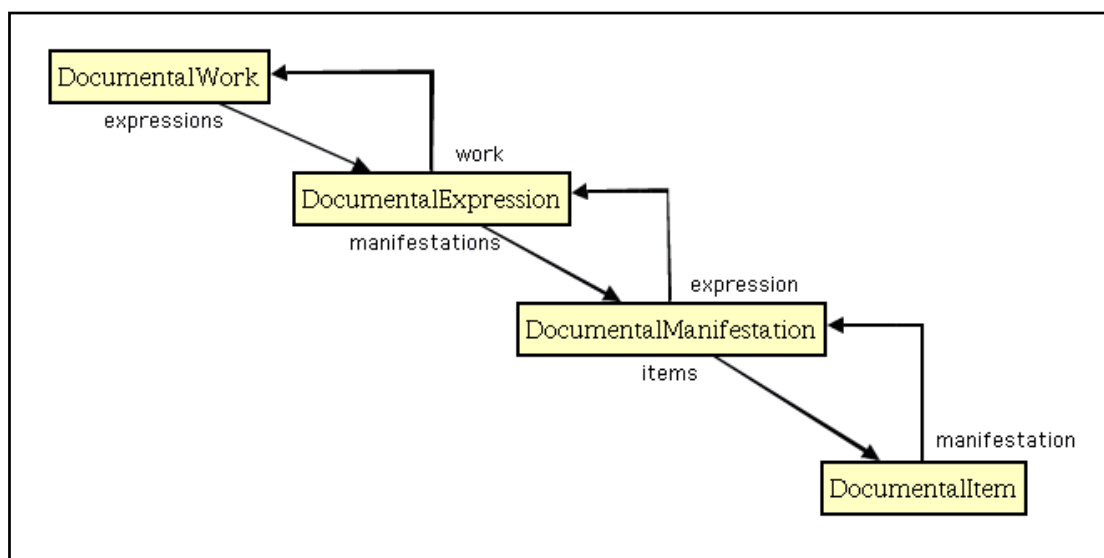
El siguiente paso es entonces hacer que el sistema establezca estos identificadores, con lo que arribamos al otro modelo: el relacional. La normalización es el punto fuerte de este modelo -el más extendido en el mundo de los programadores- dado que plantea precisamente como primera regla la no repetición de ningún dato (primer forma normal). Llevado a sus últimas consecuencias, el

modelo relacional permite una normalización total: cada entidad, por dependiente que sea, estará representada por una tabla. Para representar la doble relación "documento-autor" es necesario crear una tercera tabla "documentos-autores" que es la que contiene la relación. Pero ahora no hay una distinción entre entidades de distinto nivel: físicamente todas las entidades se corresponden con el mismo tipo de representación. Dicho de otro modo, desde el punto de vista de los archivos, no hay diferencia física entre la tabla de documentos, la tabla de autores y la tabla de la relación documentos-autores. La entidad de mayor nivel es ahora el resultado de operar dinámicamente sobre las tablas, por lo que la ecuación sería ahora **un documento = un query**. El conocimiento de la entidad a representar está por un lado y los datos por otro.

El paradigma de objetos, en tanto elimina la tradicional separación entre datos y algoritmos, permite conservar lo mejor de los dos mundos. Por un lado existen tipos de objetos de distinto nivel de abstracción y cada uno de ellos equivale a alguna entidad del dominio a representar. Pero como los objetos establecen relaciones de colaboración entre otros objetos, y estas relaciones no están en algún lugar fuera de los objetos, sino dentro de ellos mismos, entonces la rigurosidad del modelo relacional se mantiene: todo documento 'conoce' a sus autores, todo autor 'conoce' a sus documentos. La ecuación anterior se puede formular ahora como **un documento = un objeto**.

El modelo OPUS

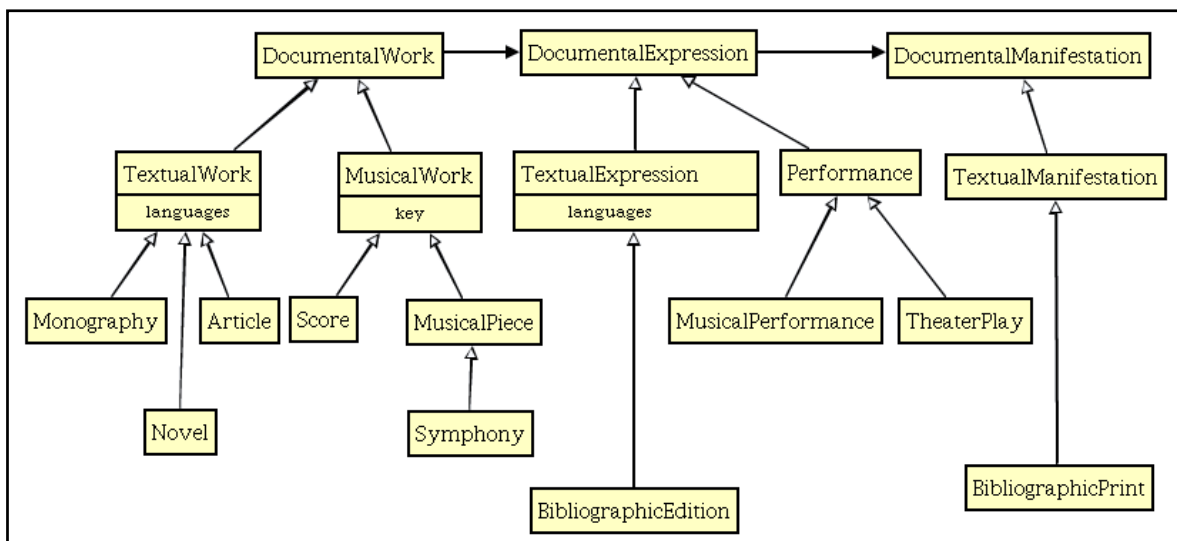
Los primeros artículos sobre el modelo FRBR (Noerr 1998, Velucci 1997) planteaban un modelo de herencia entre las entidades del grupo 1 que ha sido criticado por varios autores (Carsen 2003, Renear 2006). En estas ideas iniciales, el uso del término "herencia" no está tomado desde el paradigma de objetos y su sentido no siempre es claro. Cabe aclarar que en el reciente FRBR-₀₀ la idea de herencia es mucho más estricta y delimitada.



En nuestro modelo, las cuatro entidades del grupo 1 de FRBR-_{ER} se corresponden con cuatro clases denominadas *DocumentalWork*, *DocumentalExpression*, *DocumentalManifestation* y

DocumentalItem.

La clase *DocumentalWork* conoce a todas sus *expresiones*, mientras que cada *DocumentalExpression* conoce a su *obra*. Esta doble referencia se repite en los demás niveles, de manera que no importa cuál sea el punto de acceso, siempre se puede recorrer todos los niveles. Claramente se ve que la relación entre las entidades es de composición y no de herencia. En el informe final de FRBR-ER, si bien ya no se habla de "herencia", las entidades principales adolecen de una falta de desagregación, como se ha observado varias veces, en algunos casos al hablar de falta de "tipos" ("Tal vez a FRBR le falta un atributo "tipo" para cada una de las entidades

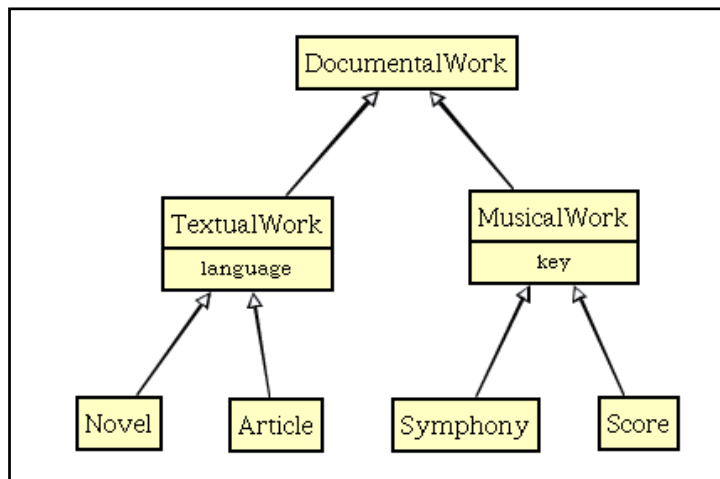


superiores" [Le Boeuf 2003b, p 11]), otras veces mencionando específicamente la herencia de clases ("La clase completa de Publicaciones –*manifestaciones*, en terminología FRBR– puede ser dividida en distintos medios: libros, films, etc." [Heaney 1995, p 141]).

Este problema se debe en parte que es un modelo entidad-relación que, al ser trasladado a una base de datos relacional, deriva en una complejidad inmanejable, y en parte también a la dependencia de las prácticas tradicionales de catalogación, que apuntan más a la descripción de las entidades mediante asignación de cadenas de caracteres que a definir las entidades en sí mismas y en su relación con el contexto. Casi todos los atributos de las entidades del grupo 1 definidas en FRBR-ER denuncian este problema. Algunos ejemplos: el atributo *Form of Work* al que se define por extensión dando los ejemplos novela, obra de teatro, poesía, ensayo, sinfonía, etc., (IFLA 1997) no tienen ningún sentido en un diseño de objetos, pues es imposible establecer algún comportamiento propio de –digamos– una sinfonía, dado que estaríamos predicando sobre una entidad mucho más abarcativa, la *obra*. La pregunta "¿Qué instrumentos participan en la obra?" podría ser dirigida a una novela, en tanto el atributo *Original Language of the Work* nos podría llevar a preguntarle "¿En qué idioma está?" a una sonata.

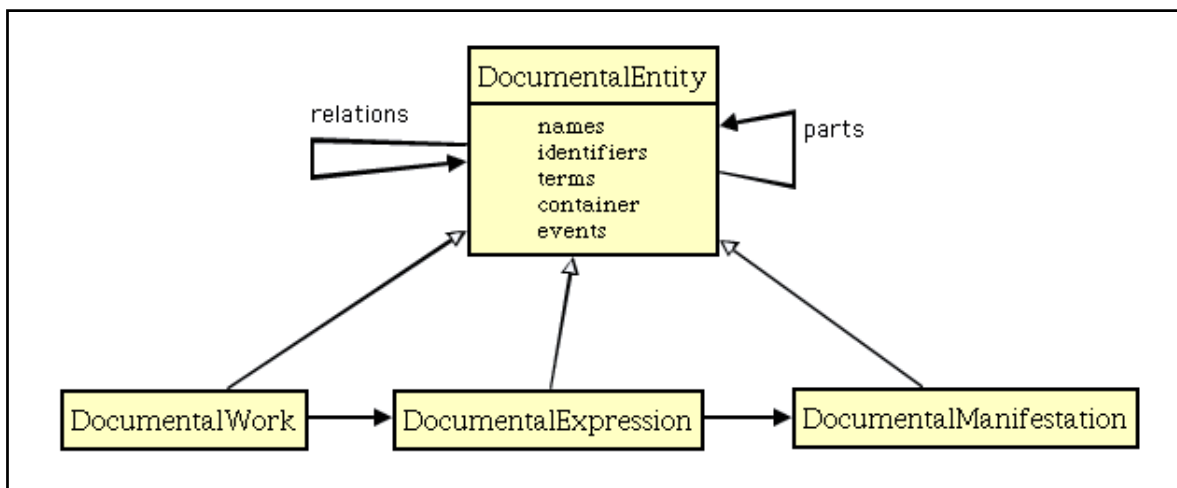
En el modelo FRBR-OO este problema no está tratado, pues el trabajo, que aún está en un estado de elaboración, se ha detenido antes. No se ha definido ninguna subclase de las entidades del grupo 1 que ejemplifique "tipos".

En nuestro modelo, la clase *DocumentalWork* es una clase abstracta que tiene el comportamiento común a toda obra y se subclasifica en clases concretas para cada tipo en particular, proceso que termina conformando los denominados *tipos documentales*. Cada una de las subclases contempla atributos y comportamiento propios. Por ejemplo, el atributo *language* sólo tendrá sentido en subclases de *DocumentalWork* cuyas expresiones contengan texto (*TextualWork*), mientras que el atributo *key* (*tonalidad*), pensado exclusivamente para el repertorio de la tradición clásico-romántica occidental, tendrá sentido en objetos que formen parte de la subclase *MusicalWork*, etc.



Este mismo proceso de subclasificación se puede continuar en los otros niveles, en la medida en que tenga sentido en cada caso. Por ejemplo, el atributo *medium of performance* en *Expresión*, tendrá sólo sentido en una subclase de *DocumentalExpression* que conceptualice las artes temporales.

Otras distinciones muy comunes se pueden diseñar siguiendo este proceso en las otras entidades del grupo 1. Pongamos un ejemplo más: la tradicional diferenciación entre edición bibliográfica e impresión o reimpresión se definen por un lado con la subclase de *DocumentalExpression* *BibliographicEdition* y por otro con la subclase de *DocumentalManifestation*, *BibliographicPrint*. Cabe aclarar que en todos estos casos, cuando hablamos de atributos no estamos hablando de cadenas de caracteres, ni de punteros a tablas, sino de objetos; esto quiere decir que cuando hablamos, por ejemplo, del atributo *language*, queremos decir que el objeto en cuestión –digamos, una edición bibliográfica- conoce a un objeto de la clase *Language* que posee el comportamiento asociado al concepto de idioma. Por razones de legibilidad estas relaciones no están representadas en los gráficos.



El problema inverso al de la subclasificación tampoco ha sido tenido en cuenta: la falta de superclases abstractas que engloben comportamientos comunes. En el proceso de diseño con objetos esta necesidad no siempre surge en los primeros momentos, cuando el objetivo es representar las entidades concretas que se evidencian a primera vista. Sin embargo, una vez que están establecidos los comportamientos de las entidades más visibles, surgen repeticiones de atributos y comportamientos comunes, que indican la necesidad de una abstracción más general que simplifique el diseño y elimine las reiteraciones.

Si bien el modelo CRM es pródigo en superclases abstractas, no puede decirse lo mismo de FRBR-₀₀.

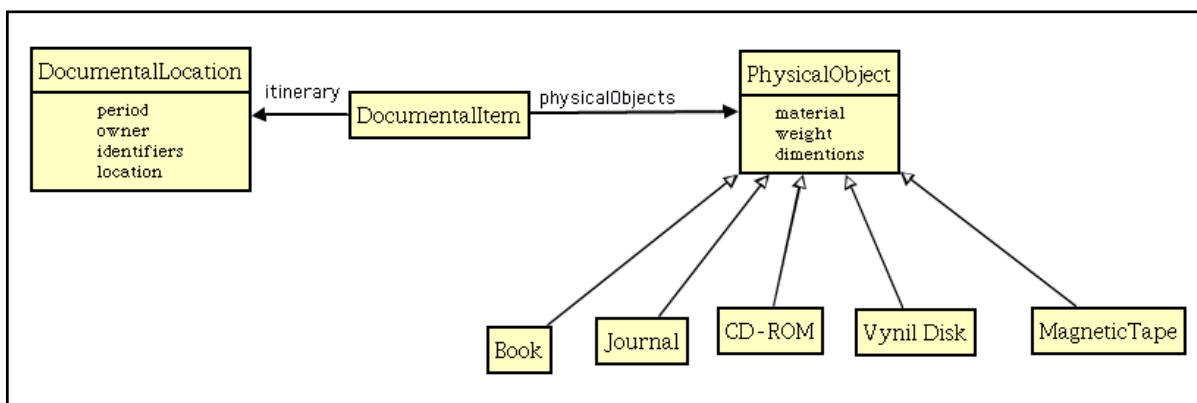
Al menos en los tres primeros niveles del grupo de entidades 1 podemos observar ciertos elementos presentes en todos:

- 1) Nombres: el título uniforme en la obra, el título en la expresión (esto es, el título que el autor quiso poner que no necesariamente es el título final), el título propio en la manifestación.
- 2) Identificadores: eventuales identificadores universales de obras y de expresiones que muy probablemente se establecerán en un futuro cercano, los ISSN, ISBN, etc., en las manifestaciones.
- 3) Relación todo-parte: una obra multi-parte, su correspondiente expresión multi-parte, agrupaciones de obras en una manifestación, como las antologías, etc.

4) Relaciones con otras entidades: todas las relaciones entre las entidades del grupo 1 que han sido detalladas en el informe final de IFLA FRBR (IFLA 1997) y en otros trabajos (Velucci 1997). Esto amerita la existencia de una superclase abstracta *DocumentalEntity* que represente al documento antes de pertenecer a un nivel determinado. Obsérvese en el gráfico que los objetos de la clase *DocumentalEntity* establecen relaciones de todo-parte y otro tipo de relaciones con otros objetos de la misma clase, y por consiguiente, con cualquiera de sus subclases. Esto quiere decir que tanto las relaciones entre los mismo niveles, (*obra-obra*, *obra-expresión*, etc.) como las combinaciones de relaciones entre niveles distintos (*obra-expresión*, *expresión-manifestación*, etc.) son soportados. En el modelo completo estas relaciones están claramente delineadas mediante los mismos procesos de diseño mostrados aquí, pero no nos es posible extendernos tanto.

El problema del *ítem* y el soporte físico

Otro problema que hemos considerado insuficientemente tratado por FRBR es el del estatus ontológico del *ítem* y la relación de éste con el soporte físico. Tanto FRBR-_{ER} como FRBR-_{OO} consideran al objeto físico como parte del mundo documental. Por ejemplo, el informe de FRBR-_{OO}



en la definición de la clase *F10 Object* dice: "Esta clase comprende ítems de naturaleza material que son unidades para la documentación", mientras que la definición de la clase *F5 Item* es: "Esta clase comprende objetos físicos (...) que portan una *F41 Publication Expression* ..." (IWG 2006, CIDOC 2004).

Nosotros, en cambio, vemos el *ítem* como una nueva entidad que surge de la acción del catalogador sobre un objeto físico al incorporarlo a una colección documental.

Ciertas propiedades del *ítem*, como por ejemplo su itinerario entre distintas colecciones -lugares en donde estuvo alojado, las fechas de egreso e ingreso, las instituciones o personas que lo poseyeron-, no son propiedades de los objetos físicos, sino sólo de éstos en tanto documentos. Si bien este itinerario impacta en el objeto físico en la forma de sellos, anotaciones, etc., cada una de estas marcas habla más del estatus del objeto en tanto documento coleccionado, clasificado e inventariado que de su historia física.

Por otro lado, no todo objeto físico que esté representado en un sistema documental deberá ser necesariamente un documento. Por poner un ejemplo que sólo una mirada ingenua puede considerar remoto: en un sistema de biblioteca virtual que represente a una biblioteca en su dimensión física, el mobiliario, los estantes, las habitaciones, etc, no son documentos, pero sí deben estar representados y poseen los mismos atributos en tanto objetos físicos que los libros o CD-ROMs.

Al quedar claro que se trata de entidades que están, al menos en parte, afuera del dominio estricto de la documentación, no es aconsejable utilizar la subclasificación: entidades como "libro", "disco", "revista" forman parte también del mundo físico y comparten en tanto objetos físicos muchas propiedades aunque desde el punto de vista documental puedan diferir enormemente; las modificaciones que sufren en su interacción con el universo físico están sujetas a las mismas leyes. Por consiguiente, en nuestro modelo definimos dos clases distintas, *DocumentalItem* y *PhysicalObject* que tienen una relación de composición, no de herencia.

Establecer identidad entre objeto físico e ítem ocasiona a otro problema más: no todo ítem se

corresponde con un único objeto físico. Por ejemplo, una grabación musical en CD-ROM representa, desde el punto de vista de una colección documental, un único ítem. No obstante, como objeto físico se compone del disco óptico, el cuadernillo y la caja de plástico, todos ellos objetos físicos con propiedades distintas, que pueden sufrir distintas alteraciones. Sólo si falta o está dañado el primero la obra deja de ser accesible; si falta el segundo hay pérdida de información referencial. Si sólo falta la caja de plástico, desde el punto de vista informativo el daño es despreciable: hasta puede reemplazarse fácilmente.

Este esquema deja afuera una distinción muy importante que está claramente establecida en la relación *manifestación-ítem*: la de ser molde y copia. Efectivamente los objetos físicos pueden tener esa relación, siendo el molde una versión "ideal" de la que luego surgen las copias. Por consiguiente, lo correcto sería que cada *manifestación* tuviese una relación con ese objeto físico "ideal", llamado aquí *PhysicalObject*, y cada *ítem* por un objeto físico concreto, *PhysicalObjectCopy* que a su vez es copia del objeto físico molde.

Las propiedades relativas a todas las copias (peso, dimensiones, material, etc.) figuran en *PhysicalObject*, mientras que las modificaciones particulares que ha sufrido cada copia, configuran el atributo *state* del objeto *PhysicalObjectCopy*.

Dejamos para otra ocasión la consideración de los documentos electrónicos, que, si bien comparten algunas características de los objetos físicos, no pueden ser considerados como tales.

El problema de la responsabilidad y la función

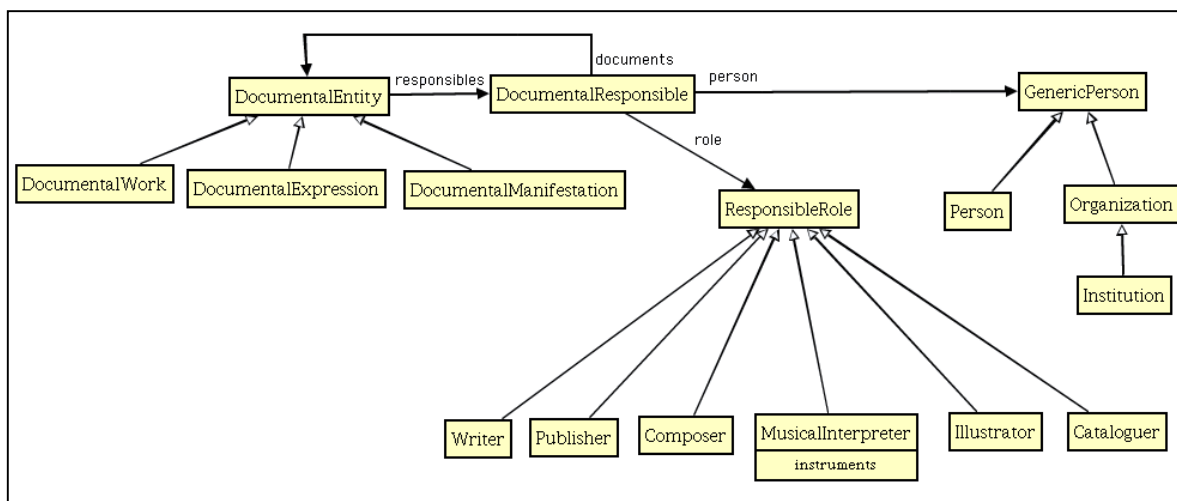
Algo semejante ocurre con el problema de la responsabilidad. En FRBR-_{ER} las entidades *Person* y *Corporate Body* son vistas sólo en su rol de participantes del proceso documental, siguiendo la tradición de las normas de catalogación. Este problema ha sido observado reiteradamente e incluso la propuesta de FRANAR lo hace notar al aclarar: "La entidad persona no refleja una persona existente en el mundo real, sino una intermedia entre el mundo real y el universo del catalogo" (IFLA UBCIM 2007) aunque al mantener la misma terminología, conserva aún cierta ambigüedad.

Un usuario de un sistema de información es también una persona: ¿por qué debería tener atributos como "documentos publicados". Esto nos obligaría a definir una clase *Usuario*, que compartiría casi todos los atributos de *Persona*, pero en una jerarquía totalmente independiente, lo que sería un grave error de diseño.

La propuesta FRBR-_{OO} en cambio, por tratarse de un proyecto de armonización con CRM "hereda" muchas de sus virtudes, y en este caso, rompe con el sesgo bibliotecológico al derivar las clases *F6 Family* y *F7 Corporate Body* de la clase de CRM *E21 Person*, que representan a los conceptos familia, entidad corporativa y persona independientemente de cual sea su función en el universo documental.

FRBR-_{OO} "hereda" también el esquema de eventos de CRM para tratar la cuestión de la responsabilidad documental en los distintos niveles, lo cual asegura que la inscripción del proceso documental en el tiempo esté asegurada, que en el FRBR original era una omisión que fue marcada varias veces. (Heaney 1997).

Sin embargo, el problema que surge inmediatamente al hablar de responsabilidad documental es el de las funciones: ¿cómo se consigna el hecho de que una persona sea autor, traductor,



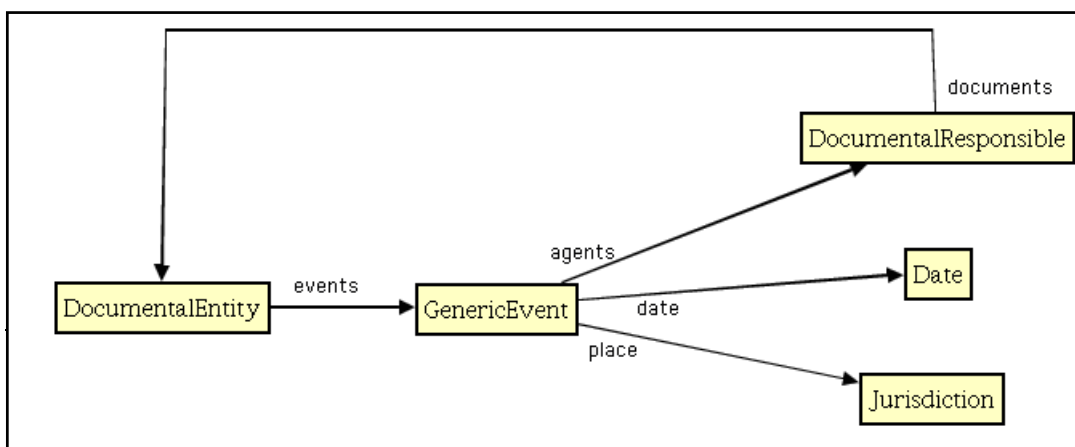
intérprete, etc., de una obra? No hay mención en las dos propuestas de FRBR a esta cuestión.

Nosotros hemos partido del esquema de eventos mencionado, pero haciendo la salvedad de que estos modelos hacen caer todo el peso del comportamiento asociado a la responsabilidad documental en los eventos (*F30 Work Conception*, *F31 Expression Creation*, *F33 Identifier Assignment*, etc). En cambio, hemos preferido poner el acento en la persona que participa de la acción y no tanto en la acción misma. Si quisiésemos representar la función "traductor" siguiendo a FRBR-00, deberíamos subclasificar *F31 Expression Creation* y arribar a una clase *Translation*. En cambio hemos preferido establecer que cada función de responsabilidad tenga una clase asociada a la persona y no a la acción; en otras palabras, preferimos una clase *Translator* a una clase *Translation*

Lo que hemos modelado es una clase que representa a las personas en tanto responsables de documentos, *DocumentalResponsible*, que delega en la persona el comportamiento asociado a la misma, pero que conoce las relaciones y atributos de esa persona en su interacción específica con documentos. Toda *DocumentalEntity*, es decir una obra, una expresión o una manifestación, conoce a su o sus responsables; a su vez todo *DocumentalResponsible* conoce a sus obras, expresiones o manifestaciones. Por medio del mecanismo de delegación, toda *DocumentalEntity* puede comunicarse con la persona en sí, por lo que esta relación no se pierde.

Esta clase permite además solucionar el problema de las distintas funciones de responsabilidad.

Ahora la función de ser autor, editor, traductor, etc. no reside en la persona, pues una persona puede cumplir diversos roles en diversos documentos. Cada *DocumentalResponsible* representa a una determinada persona, que tiene un determinado rol en determinados documentos.

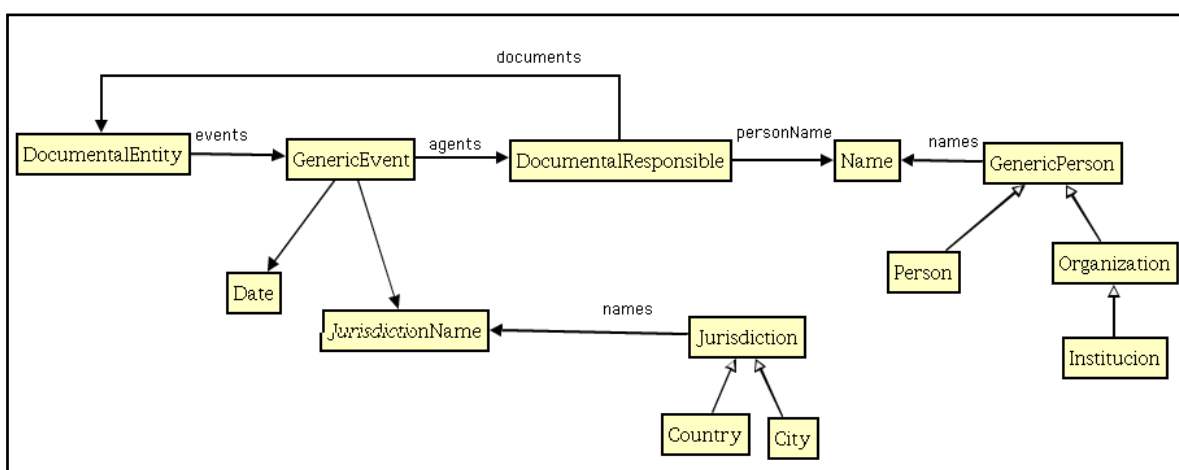


Con esta aproximación es posible definir características propias de una determinada función de responsable sin necesidad de recurrir a atributos extemporáneos. Por ejemplo, para consignar el o los instrumentos que ejecuta un músico en una determinada obra, la clase *MusicalInterpreter* puede definir el atributo *instruments*.

Como mencionamos recién, hemos adoptado el modelo de eventos de CRM y FRBR-₀₀, que usando la terminología de nuestro modelo podría resumirse así: a toda *DocumentalEntity* le suceden una serie de eventos, que consisten en acciones realizadas sobre el documento en un lugar determinado (*Jurisdiction*), una fecha determinada (*Date*) y realizadas por un agente (un *DocumentalResponsible*). El modelo contempla diversas subclases de eventos pero no nos es posible extendernos más.

El problema de los nombres

Otro problema relacionado con la responsabilidad, pero que abarca casi la totalidad de las entidades del mundo documental es el de los nombres, que ya mencionamos en nuestro trabajo anterior (Carsen 2003). No es posible ahora detenernos en cómo modelan CRM y FRBR-₀₀ los nombres y cuál es el detalle de nuestra propuesta. No obstante queremos destacar un aspecto



importante del modelo para dar una idea clara de cómo se establece el esquema final con respecto al problema de la responsabilidad. La relación que establece la *DocumentalEntity* con el *DocumentalResponsible* no es directa, sino que está mediada por el nombre. Esto significa que un documento tiene como responsable uno de los varios *nombres* que una persona puede tener. De esta forma se mantiene tanto la inscripción original del documento, es decir la *mención de responsabilidad*, como la relación con la persona en sí (usando el mecanismo de delegación ya referido). Esto soluciona varios problemas asociados a los nombres personales sin producir repeticiones: el problema de los seudónimos, de las variantes de nombres, los apellidos de casada, etc. El mismo método se utiliza al relacionar los eventos con los lugares: el lugar de edición de un documento es uno de los nombres que tiene (o que ha tenido) una determinada jurisdicción (una ciudad en este caso).

El problema de las series

Un aspecto especialmente crítico de nuestra tarea ha sido la representación de series, dado que una de las colecciones de las que debemos ocuparnos es la del ISSN nacional. La ubicación de las series en FRBR ha tenido poco tratamiento hasta ahora en la literatura específica, y en algunos casos no hay acuerdo en cuanto a la aplicación del nivel *obra* a la serie (Berg 2004).

Nuestro punto de partida ha sido lo enunciado por Pat Riva (Riva 2003) con respecto a la relación entre las series y las entidades de FRBR: La *obra* es la serie subyacente considerada como un todo, las *expresiones* abarcan el texto en el idioma original, las diversas traducciones, sus versiones en audio, etc., las *manifestaciones* consisten en la impresión original, sus reimpressiones, sus producciones en microfilm, en CD y sus versiones remotas, mientras que el *ítem* lo constituye la suscripción completa. Esta opinión parece ser compartida, entre otros, por Yee cuando afirma que "los usuarios seguramente considerarán tanto la base de datos (las versiones electrónicas de la serie) como la revista que buscan, ambas como versiones diferentes de la misma obra" (Yee 2003).

Hemos dejado de lado por el momento las polémicas con respecto a la posibilidad de un nivel *superwork* que muchos ven apropiado para las series, en la creencia de que los problemas que han hecho concebir este nivel "extra" pueden ser solucionados mediante el establecimiento de relaciones adecuadas entre los cuatro niveles básicos, siguiendo la idea de *familia de la serie* (Riva 2003)

Entre los muchos problemas que suscita la modelización de las series y del material bibliográfico en general, la cuestión de los congresos tiene un lugar especial. Contra la tradición de pensar los congresos como entidades documentales, FRBR-00 ha distinguido claramente que los congresos son eventos independientes que pueden llegar a tener alguna relación con documentos, incluyéndolos como ejemplos de la clase *F11 Event*. Pero el modelo se detiene exactamente ahí: un congreso es por ahora tan sólo un ejemplo de esa clase.

En nuestro modelo hemos definido una clase *Meeting*, cuyas instancias pueden (o no) tener relación con documentos, tanto sean series como cualquier otro tipo de documento. Como subclases de *FrequentEvent* hereda los atributos *frequency* y *period*, en tanto incorpora los atributos *languages* y *reponsibles*. Obsérvese que este último atributo consiste en una colección de objetos de la clase *GenericPerson*, lo que reafirma la conveniencia de una clara distinción de entidades en el diseño; estamos reusando una clase, cosa que no podríamos haber hecho si las personas hubiesen quedado acopladas a su rol de responsables documentales. El objeto persona en este caso se refiere a el o los responsables del congreso, no a los responsables (autores, editores, etc.) de un documento.

La implementación

Para terminar, querríamos dar algunos detalles de implementación del programa Opus.

Dado que nuestro trabajo se enmarca dentro del paradigma de objetos, hemos utilizado el lenguaje de programación que mejor lo representa y que no incorpora ningún elemento de otros paradigmas: Smalltalk. Dentro de los dialectos de Smalltalk, hemos elegido Squeak por tratarse de

un programa de código abierto. Nuestro trabajo ha producido, además de la aplicación en sí, algunos *frameworks* que pondremos en breve a disposición de la comunidad informática: Atón, para interactuar desde Smalltalk con bases de datos ISIS y Smallfaces, una ampliación del framework Seaside para construcción de interfaces web.

Para mantenernos lo más posible dentro de una misma metodología de trabajo, para la persistencia de los objetos hemos optado por una base de datos de objetos. Teniendo en cuenta las limitaciones presupuestarias a que estamos sometidos, hemos optado por una base de objetos gratuita: Magma.

A esta base de datos de objetos estamos migrando el material existente en las bases de datos ISIS del CAICYT: la base de datos de ISSN nacional, la base de datos del Catálogo Colectivo de Publicaciones Periódicas y las bases de datos de las bibliotecas DOCSA, GIETZ. y REMCyTA. Actualmente está en uso la aplicación para el registro de ISSN nacional.

Bibliografía

- Berg, E. S. (2004). Implementing FRBR: A comparison of the two relational models: IFLA's FRBR model and Taniguchi's expression-prioritized model. Recuperado en Noviembre 3, 2007 de http://home.hio.no/~bagheri/Master_thesis/Implementing_FRBR.pdf
- Carsen, T. & Garcia, H. & Mabragaña, C. & Manzanos, N. (2003). Modelo de descripción documental basado en el paradigma de objetos. En José Antonio, Frías Montoya and Travieso Rodríguez, Crispulo, Eds. Proceedings VI Congreso del Capítulo Español de ISKO, Salamanca (Spain). Recuperado en Noviembre 3, 2007 de: <http://eprints.rclis.org/archive/00008585/>
- Dugan, B. (1994) *Simula and Smalltalk: A Social and Political History*. Recuperado en Octubre 2002 de <http://www.cs.washington.edu/homes/dugan/history.html>
- Grupo de estudio de la IFLA sobre los Requisitos Funcionales de los Registros Bibliográficos. (2004). *Requisitos Funcionales de los Registros Bibliográficos: Informe Final*. Madrid: Ministerio de Cultura.
- Heaney, M.(1995) *Object-oriented cataloguing*. En *Information Technology and Libraries*, vol 16, no. 3, pp.135-153.
- Heaney, M. (1997) *Time is of the essence: some thoughts occasioned by the papers contributed to the International Conference on the Principles and Future Development of AACR*. Recuperado en Noviembre 3, 2007 de <http://www.bodley.ox.ac.uk/users/mh/time978a.htm>.
- IFLA Study Group on the functional requirements for bibliographic records.(1998). *Functional requirements for bibliographic records: final report*.- Munich: K. G. Saur.
- IFLA UBCIM Working Group on Functional Requirements and Numbering of Authority Records (FRANAR).(2007). *Functional Requirements for Authority Records: A Conceptual Model*. Recuperado en Noviembre 3, 2007 de <http://www.ifla.org/VII/d4/FRANAR-ConceptualModel-2ndReview.pdf>

- International Working Group on FRBR and CIDOC CRM Harmonization. (2007). *FRBR object-oriented definition (version 0.8.1)* . Recuperado en Noviembre 3, 2007 en http://www.ifla.org/VII/s13/wgfrbr/FRBR_oo_V.0.8.1c.pdf
- Ingalls, D. (1981). *Design Principles Behind Smalltalk*. En *Byte*, vol. 6, no. 8. (pp 286-298).
- Kay, A. (1993). *The Early History of Smalltalk*. En *ACM SIGPLAN Notices*, vol 28, no3.
- Le Boeuf, P. (2003a). *The book, the bug and the bangle: a parallel and a paradox*. En *Sharing the knowledge: international CIDOC CRM Symposium*, Washington DC. Recuperado en Septiembre 9, 2005 de: http://cidoc.ics.forth.gr/docs/symposium_presentations/leboeuf_bookbugbangle_revised.doc
- Le Boeuf, P.(2003b). *Brave new FRBR world*. En First IFLA Meeting of Experts on an International Cataloguing Code, Frankfurt am Main, Germany [on line]. Frankfurt am Main: Die Deutsche Bibliothek.
- Liu, C. (1996). *Smalltalk, Objects, and Design*. New York: toExcel.
- Heaney, M. (1995). *Object-oriented cataloguing*. *Information Technology and Libraries*, vol 16, no. 3, (pp.135-153).
- Moles, A.. (1971). *Sociodinámica de la cultura*. Buenos Aires: PAIDOS.
- Mortensen, S.E. (2001). *Why Java Isn't Smalltalk: An Aesthetic Observation*. En *Smalltalk Chronicles: vol. 3*, no 1.
- Moya Anegón, F. (1995). *Los sistemas integrados de gestión bibliotecaria: Estructuras de datos y recuperación de información*. Madrid: ANABAD.
- Noerr, P. & Goossens, P. & Matei, D.& Otten, P. & Peruginelli, S & Witt, M. (1998). *User Benefits from a New Bibliographic Model: Follow-up of the IFLA Functional Requirements Study*. Recuperado en Noviembre 3, 2007 de <http://www.ifla.org/IV/ifla64/084-126e.htm>
- Renear, A. & Choi, Y. (2006). *Modeling Our Understanding, Understanding Our Models: The Case of Inheritance in FRBR*. End Grove, Andrew (Eds). *Proceedings 69th Annual Meeting of the American Society for Information Science and Technology (ASIST) 43*, Austin (US). Recuperado en Noviembre 3, 2007 de <http://eprints.rclis.org/archive/00008158/>
- Riva, P. (2003). *Defining the Boundaries : FRBR, AACR and the Serial*. En *Serials Librarian*. Vol. 45. No. 3. (pp.15-21).
- Taylor, A. D. (1997). *Object Technology*. [S.L.]: Addison Wesley.
- Vellucci, S. L. (1997). *Bibliographic Relationships*. En *International Conference on the Principles and Future Development of AACR*. Toronto: University of Toronto Faculty of Information Studies.
- Xerox Palo Alto Research Center - Learning Research Group. (1976) *Personal Dynamic Media*. Palo Alto: Xerox Corporation
- Yee, M. (2000). *Lubetzky's work principle*. En Tschera Harkness Connell & Robert L. Maxwell, (eds), *The future of cataloging: insights from the Lubetzky Symposium*. [S. l.] : cop. the American Library Association.